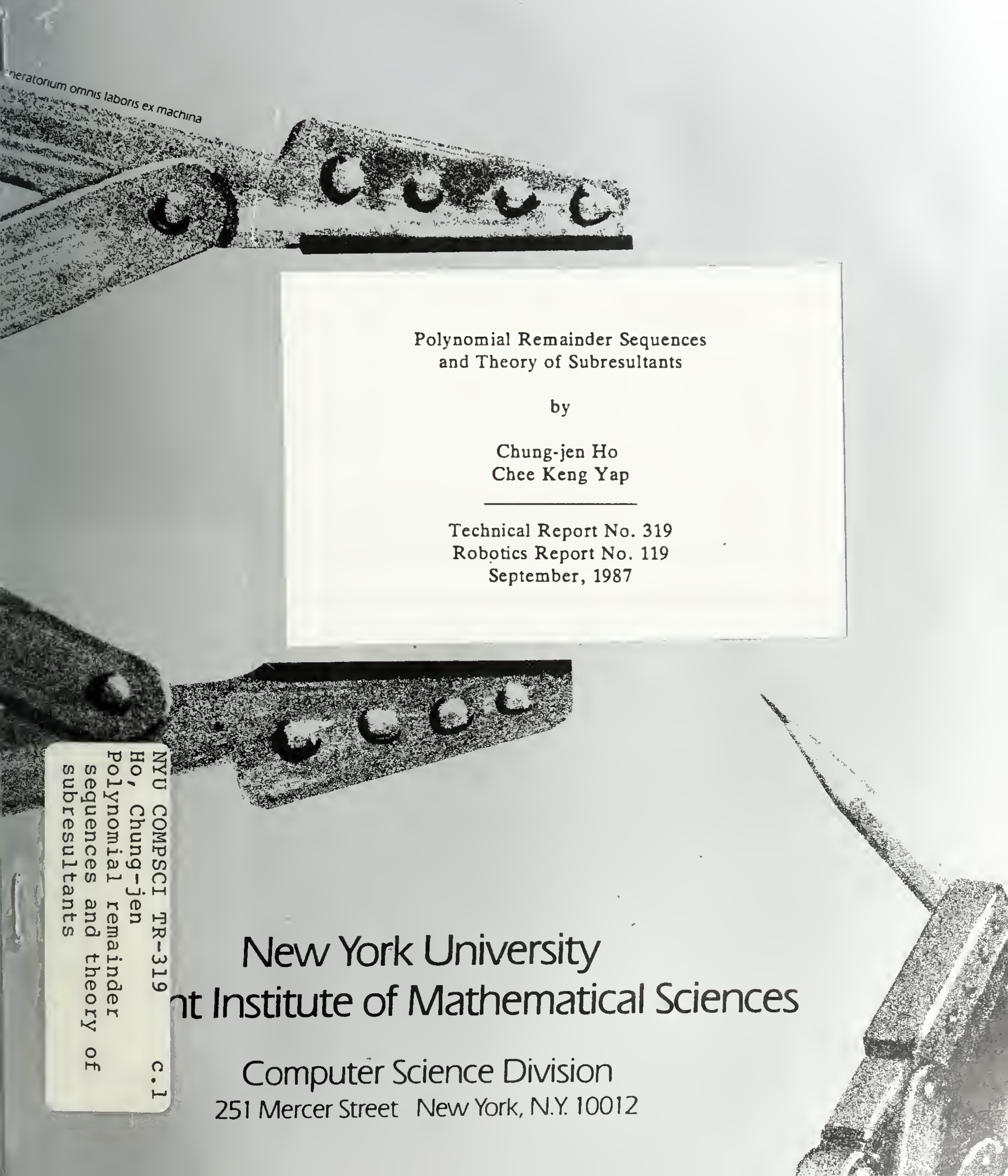
Robotics Research Technical Report



Polynomial Remainder Sequences
and Theory of Subresultants

by

Chung-jen Ho
Chee Keng Yap

Technical Report No. 319
Robotics Report No. 119
September, 1987

NYU COMPSCI TR-319 c.1
Ho, Chung-jen
Polynomial remainder
sequences and theory of
subresultants

New York University
nt Institute of Mathematical Sciences

Computer Science Division
251 Mercer Street New York, N.Y. 10012

Polynomial Remainder Sequences
and Theory of Subresultants

by

Chung-jen Ho
Chee Keng Yap

Technical Report No. 319
Robotics Report No. 119
September, 1987

New York University
Dept. of Computer Science
Courant Institute of Mathematical Sciences
251 Mercer Street
New York, New York 10012

Work on this paper has been supported by NSF grants DCR-84-01898 and DCR-84-01633.

Polynomial Remainder Sequences and Theory of Subresultants

Chung-jen Ho and Chee Keng Yap [*]

Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, NY 10012

August 31, 1987

1 Introduction

There are many basic applications where we want to compute the GCD of a pair of polynomials, together with the associated sequence of remainders. These include Diophantine equations, Sturm sequences, elimination theory, algebraic cell decompositions [Y85]. There is an obvious way to apply Euclid's algorithm for integers to a pair $p(x)$, $q(x)$ of polynomials in $F[x]$ for any field F: compute the sequence

$$p_0, p_1, \ldots, p_h, p_{h+1} \tag{1}$$

where $p_0 = p$, $p_1 = q$, $p_{h+1} = 0$ and p_{i+1} is the remainder of p_{i-1} divided by p_i,

$$p_{i+1} = \mathrm{rem}(p_{i-1}, p_i), \quad i = 1, 2, \ldots, h.$$

---

[*] This work is supported by NSF grants DCR-84-01898 and DCR-84-01633.

Example 1

$$\begin{aligned} p_0 &= x^5 + 2x^4 + 3x^2 - x + 2 \\ p_1 &= 3x^3 - x + 2 \\ p_2 &= 3x^2 - \frac{20x}{9} + \frac{16}{9} \\ p_3 &= -\frac{275x}{243} + \frac{166}{243} \\ p_4 &= \frac{115668}{75625} \\ p_5 &= 0 \end{aligned}$$

However, if $p(x), q(x) \in I[x]$ where I is an integral domain, the remainder is in general not well-defined. To handle this, one defines the *pseudo-remainder* of $p(x)$ divided by $p'(x)$ to be the following polynomial $\mathbf{prem}(p, p')$: if $\mathbf{deg}(p) < \mathbf{deg}(p')$ or $p' = 0$ then $\mathbf{prem}(p, p') = p$. Otherwise, $\mathbf{prem}(p, p')$ is the polynomial $r(x) \in I[x]$ such that there exists $q(x) \in I[x]$,

$$r(x) + p'(x)q(x) = b^{d+1}p(x), \qquad \mathbf{deg}(r) < \mathbf{deg}(p') \tag{2}$$

where $b = \mathbf{lead}(p'(x))$ is the leading coefficient of p' and $d = \mathbf{deg}(p(x)) - \mathbf{deg}(p'(x))$. Thus by definition,

$$\mathbf{prem}(p, p') = \mathbf{rem}(b^{d+1}p, p').$$

Exercise 1 Use long division to compute $\mathbf{prem}(p, p')$ where $p(x) = x^6 + 2x^4 + 3x^2 - x + 2$ and $p'(x) = 3x^3 - x + 2$.

Exercise 2 Show that $\mathbf{prem}(p, p')$ is well-defined for all $p, p' \in I[x]$.

Definition. Let $p(x), q(x) \in I[x]$. Then p and q are *similar*, denoted $p \sim q$, if there exists $a, b \in I$ such that

$$ap(x) = bq(x)$$

A special case of this is where a and b are units, in which case p and q are said to be *associates* of each other.

Definition. Let $p(x), q(x)$ be non-zero polynomials in $I[x]$ with $\mathbf{deg}(p) \geq \mathbf{deg}(q)$. Then a sequence $(p_0, p_1, \ldots, p_h)$ $(h \geq 1)$ is called a *polynomial remainder sequence* (PRS) of p and q if $p_0 = p$, $p_1 = q$ and for $i = 2, 3, \ldots, h$,

$$p_i \sim \mathbf{prem}(p_{i-2}, p_{i-1})$$

and

$$\mathbf{prem}(p_{h-1}, p_h) = 0.$$

Following Collins [C66], we call a PRS $(p_0, p_1, \ldots, p_h)$ the *Euclidean PRS* if $p_i = \mathbf{prem}(p_{i-2}, p_{i-1})$ for $i = 2, \ldots, h$. The difficulty with the Euclidean PRS is its unmanageable coefficient growth. Thus the main issue in this area of research is to discover efficient methods of computing polynomial remainder sequences.

Example 2 (Knuth's example) Displaying only coefficients, the following shows an Euclidean PRS:

x^8	x^7	x^6	x^5	x^4	x^3	x^2	x	1
1	0	1	0	−3	−3	8	2	−5
		3	0	5	0	−4	−9	21
				−15	0	3	0	−9
						15795	30375	−59535
							1254542875143750	−1654608338437500
								12593338795500743100931141992187500

Remarks. This report is extracted from our lecture notes [Y85] (which accounts for the presence of "Exercises"). It, in turn, was an attempt to expound the theory of subresultants along the lines suggested by Loos [BCL83]. We decided to publish these notes in view of the fundamental importance of subresultants and the fact that the theory is not easily accessible. Furthermore, as we shall point out, these were some major oversights in Loos [BCL83] and the correct version is intricate enough to deserve careful accounting. The Loos-Habicht approach in [BCL83] is basically different from the well-known account of Collins [C67] and Brown and Traub [BT71]. We feel that in some ways, this alternative approach is more informative and "digestible".

2 The Primitive PRS

Let I be an Euclidean domain. A polynomial $p \in I[x]$ is said to be *primitive* if its coefficients are relatively prime. It is easily seen that we can express any polynomial p in the form

$$p(x) = c \cdot q(x) , \quad c \in I \tag{3}$$

where $q(x)$ is primitive. For instance, over $\mathbf{Z}[x]$, $p(x) = -2x^2 - 4x + 12$ is not primitive, and we can write $p(x) = c \cdot (x^2 + 2x - 6), c = -2$.

Lemma 1 *The factorization (3) is unique in the sense that if $p(x) = c \cdot q(x) = c' \cdot q'(x)$ then*

$$c = uc' , \quad uq(x) = q'(x)$$

where u is a unit in I. In other words, c and $q(x)$ are unique up to associates.

Proof. Exercise. **Q.E.D.**

In order to get uniqueness in this lemma, it is conventional to choose a distinguished element among each equivalence class of associates in I. Then we define the *content* of $p(x)$ to be the element $\mathbf{cont}(p(x)) \in I$ and the *primitive part* of $p(x)$ to be the primitive polynomial $\mathbf{prim}(p(x))$ such that

$$p = \mathbf{cont}(p) \cdot \mathbf{prim}(p)$$

and $\mathbf{lead}(\mathbf{prim}(p))$ is distinguished among its associates. A basic result about primitive polynomials is the following.

Lemma 2 (*Gauss's Lemma*) *Let I be a unique factorization domain. If $p(x), q(x) \in I[x]$ are primitive, so is their product $p(x)q(x)$.*

Proof. Exercise. **Q.E.D.**

Exercise 3 Prove that if I is a unique factorization domain, so is $I[x]$.

Exercise 4 Show that for $p, q \in I[x]$, $\mathbf{cont}(\mathrm{GCD}(p,q))$ and $\mathrm{GCD}(\mathbf{cont}(p), \mathbf{cont}(q))$ are associates, and $\mathbf{prim}(\mathrm{GCD}(p,q))$ and $\mathrm{GCD}(\mathbf{prim}(p), \mathbf{prim}(q))$ are associates.

We now define another natural family of PRS's. A PRS $(p_0, \ldots, p_k)$ is a *Primitive PRS* if for each $i = 1, \ldots, k-1$,

$$p_{i+1} = \frac{\mathbf{prem}(p_{i-1}, p_i)}{\mathbf{cont}(\mathbf{prem}(p_{i-1}, p_i))} = \mathbf{prim}(\mathbf{prem}(p_{i-1}, p_i)).$$

Exercise 5 Compute the primitive PRS for Knuth's example above. (Check: the final constant polynomial in the primitive PRS should be 143193869.)

Clearly, this PRS has the minimal coefficient growth among all PRS. The need to compute the content of a polynomial (= the GCD of all the coefficients of the polynomial) could present difficulty. The coefficient growth is only linear in the univariate case and this may be a reasonable algorithm if computing GCDs over the coefficient domain I is inexpensive. However, in the case of multivariate polynomials, this assumption is no longer true.

3 The Subresultant PRS

The primitive PRS is an example of a family of PRS which has the following form: Let

$$\bar{\beta} = (\beta_1, \beta_2, \ldots, \beta_{h-1}), \quad h \geq 1 \tag{4}$$

be a (possibly empty) sequence of elements in I. A PRS $(p_0, \ldots, p_h)$ is said to be *based on* the sequence (4) if for each $i = 2, \ldots, h$,

$$p_i = \frac{\mathbf{prem}(p_{i-2}, p_{i-1})}{\beta_{i-1}}.$$

By specifying different sequences of $\bar{\beta}$, we get different families of PRS. The computational properties of the corresponding PRS algorithm depends on how easily the β_i's can be computed. We now present what is essentially the best PRS algorithm of this kind. It was discovered by Collins [C67] around 1967 with the present form due to Brown [B71].

Definition. Let $(p_0, p_1, \ldots, p_h)$ be a PRS. For each i, let $\delta_i = \deg(p_i) - \deg(p_{i+1})$ and $c_i = \mathrm{lead}(p_i)$. Then the PRS is called a *subresultant PRS* if it based on the following $\bar{\beta}$ sequence:

$$\beta_1 = (-1)^{\delta_0 + 1}$$

and for $i = 1, \ldots, h-2$,

$$\beta_{i+1} = (-1)^{1+\delta_i} (\psi_i)^{\delta_i} c_i \tag{5}$$

where the auxiliary sequence $(\psi_1, \psi_2, \ldots, \psi_h)$ is defined by

$$\psi_0 = 1$$

$$\psi_i = \psi_{i-1} \left(\frac{c_i}{\psi_{i-1}} \right)^{\delta_{i-1}} = \frac{c_i^{\delta_{i-1}}}{(\psi_{i-1})^{\delta_{i-1}-1}}, \qquad i = 1, \ldots, h \tag{6}$$

It is quite easy to implement the subresultant PRS algorithm. On the other hand, it is not obvious why the algorithm works. More specifically, we have to show that each β_i is a divisor of (the content of) $\mathrm{prem}(p_{i-1}, p_{i-2})$. Furthermore, it turns out that the ψ_i's are elements of the coefficient domain I, in fact , they are determinants of certain coefficients of p_0 and p_1; this fact was not known in the early papers on the subresultant PRS algorithm. These facts about β_i and ψ_i will be proved in Section 7.

Exercise 6 Produce an algorithm that implements the subresultant PRS algorithm. You may use the fact that the ψ_i's are in I.

We next investigate the theory of subresultants which will reveal the structure of PRS's in general and of the subresultant PRS's in particular.

4 Subresultants

One of the first to realize that polynomial remainder sequences are intimately connected with determinants was Collins [C66]. Our notations follow Loos [BCL83]. We begin by showing how pseudo-remainders can be expressed using determinants. For any square matrix M, we let $\det(M)$ denote its determinant.

Definition. Let M be a $m \times n$ matrix where $m \leq n$. Then the *determinant polynomial* of M, $\mathrm{detpol}(M)$, is

$$\det(M_m)x^{n-m} + \det(M_{m+1})x^{n-m-1} +, \ldots, + \det(M_{n-1})x + \det(M_n)$$

where M_i $(i \geq m)$ is the square submatrix of M consisting of the first $m-1$ columns and ith column of M. We call $\det(M_m)$ the *nominal leading coefficient* of $\mathrm{detpol}(M)$.

Note that the (actual) degree of $\mathrm{detpol}(M)$ is at most $n-m$ since the nominal leading coefficient may vanish. Next we introduce a convenient notation. Let $A_i(x) \in I[x]$ $(i = 1, \ldots, m)$ be polynomials and let $n = 1 + \max_{1 \leq i \leq m} \deg(A_i)$. Then $\mathrm{mat}(A_1, A_2, \ldots, A_m)$ denotes the $m \times n$ matrix whose $(i, n-j)$th entry is the coefficient of x^j in $A_i(x)$. We shall write $\mathrm{detpol}(A_1, A_2, \ldots, A_m)$ as short hand for $\mathrm{detpol}(\mathrm{mat}(A_1, \ldots, A_m))$. The basic connection between determinants and polynomial division is revealed in the following lemma.

Lemma 3 *Let* $A, B \in I[x]$, $\deg(A) = m \geq n = \deg(B)$. *Then*

$$\mathrm{detpol}(x^{m-n}B, x^{m-n-1}B, \ldots, xB, B, A) = \mathrm{prem}(A, B).$$

Proof. Write $B(x) = \sum_{i=0}^{n} b_i x^i$ and let $A(x) = \sum_{i=0}^{m} a_i x^i$ where $b = b_n = \mathrm{lead}(B)$. Let

$$\begin{aligned} M &= \mathrm{mat}(x^{m-n}B, x^{m-n-1}B, \ldots, xB, B, b^{m-n+1}A). \\ &= \begin{pmatrix} b_n & b_{n-1} & \cdots & \cdots & b_0 & & & \\ & b_n & \cdots & \cdots & b_1 & b_0 & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & & & b_n & \cdots & \cdots & \cdots & b_0 \\ b^{m-n+1}a_m & b^{m-n+1}a_{m-1} & \cdots & b^{m-n+1}a_n & \cdots & \cdots & \cdots & b^{m-n+1}a_0 \end{pmatrix} \end{aligned}$$

Applying Gaussian elimination that corresponds to the long-division of $b^{m-n+1}A$ by B, we eventually reduce M to

$$M' = \begin{pmatrix} b_n & b_{n-1} & \cdots & \cdots & b_0 & & & \\ & b_n & \cdots & \cdots & b_1 & b_0 & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & & & b_n & \cdots & \cdots & \cdots & b_0 \\ & & & & c_{n-1} & \cdots & \cdots & a_0 \end{pmatrix}$$

where the bottom row corresponds to $\mathbf{prem}(A, B) = \sum_{i=0}^{n-1} c_i x^i$. Here $n - 1$ is only a nominal degree of $\mathbf{prem}(A, B)$ since any of the c_j's may vanish. Clearly

$$\mathbf{detpol}(M) = \mathbf{detpol}(M')$$

by basic properties of determinants. We observe that

$$\mathbf{detpol}(M') = b^{m-n+1}\mathbf{prem}(A, B)$$

since M' is upper triangular. On the other hand,

$$\mathbf{detpol}(M) = b^{m-n+1}\mathbf{detpol}(x^{m-n}B, \ldots, B, A).$$

These three equations imply the lemma. **Q.E.D.**

Definition. Let $A, B \in I[x]$ with $\mathbf{deg}(A) = m > 0$ and $\mathbf{deg}(B) = n > 0$. For any i, $0 \leq i < \mathbf{min}\{m, n\}$, define the ith *subresultant* of A and B to be

$$\mathbf{sres}_i(A, B) = \mathbf{detpol}(x^{n-i-1}A, x^{n-i-2}A, \ldots, A, x^{m-i-1}B, x^{m-i-2}B, \ldots, B).$$

The ith *principal subresultant coefficient* of A and B, denoted $\mathbf{psc}_i(A, B)$, is defined to be the nominal leading coefficient of $\mathbf{sres}_i(A, B)$.

Observe that $\mathbf{mat}(x^{n-i-1}A, x^{n-i-2}A, \ldots, A, x^{m-i-1}B, x^{m-i-2}B, \ldots, B)$ has $m + n - 2i$ rows and $m + n - i$ columns and $\mathbf{deg}(\mathbf{sres}_i(A, B))$ is at most i. In particular, $\mathbf{sres}_0(A, B) \in I$.

For convenience, we extend the definition of subresultant $\mathbf{sres}_i(A, B)$ to the cases $\mathbf{min}\{m, n\} \leq i < \mathbf{max}\{m, n\} - 1$:

$$\mathbf{sres}_i(A, B) = 0 \quad \text{for } \mathbf{min}\{m, n\} < i < \mathbf{max}\{m, n\} - 1$$

$$\mathbf{sres}_l(A,B) = \mathbf{lead}(X)^{|m-n|-1}X \quad \cdot$$

where $l = \min\{m,n\}$ and $X = A$ if $m > n$; B if $n > m$.

The sequence

$$(S_{w+1}, S_w, \ldots, S_1, S_0)$$

where $w+1 = \max\{m,n\}$, $S_{w+1} = A$, $S_w = B$ and $S_i = \mathbf{sres}_i(A,B)$ for $i = w-1, \ldots, 1, 0$ is called the *subresultant chain* of A and B. It is convenient to extend the definition of principal subresultant coefficient $\mathbf{psc}_i(A,B)$ to the cases $l \leq i \leq w+1$:

$$\mathbf{psc}_i(A,B) = \mathbf{lead}(S_i) = 0 \quad \text{for } l \leq i < w$$

$$\mathbf{psc}_w(A,B) = \mathbf{lead}(S_w) = \mathbf{lead}(B) \quad \text{and} \quad \mathbf{psc}_{w+1}(A,B) = 1.$$

Note that $\mathbf{psc}_{w+1}(A,B)$ is not defined as $\mathbf{lead}(A) = \mathbf{lead}(S_{w+1}(A,B))$ as one might have expected.

We remark that S_0 is the classical "resultant" of A and B and it is well-known that $S_0 = 0$ precisely when A and B have a common non-constant factor, i.e., when $\deg(\mathrm{GCD}(A,B)) > 0$. The subresultants generalizes the classical resultant in that if $S_0, S_1, \ldots, S_{i-1}$ all vanish but S_i does not then $\deg(\mathrm{GCD}(A,B)) = i$. These remarks will be consequences of theorems to be shown.

Exercise 7 1. Let $\deg(A(x)) = m, \deg(B(x)) = n$ and $m \geq n > k$. Show that $\deg(\mathrm{GCD}(A,B)) \geq k+1$ iff there exists polynomials $U_k(x)$ and $V_k(x)$, not both zero, $\deg(U_k) \leq n-k-1$ and $\deg(V_k) \leq m-k-1$, such that

$$A(x)U_k(x) - B(x)V_k(x) = 0. \tag{7}$$

Hint: use the fact that A, B are in a unique factorization domain.

2. Suppose that we already know that $\deg(\mathrm{GCD}(A,B)) \geq k$. Show that if there exists U_k, V_k as above, and $C_k(x)$ of degree at most $k-1$ such that

$$A(x)U_k(x) - B(x)V_k(x) = C(x) \tag{8}$$

then (7) has a non-trivial solution. (The reverse implication is trivial.)

3. Show that (8) is equivalent to $\mathbf{psc}_k(A, B) = 0$. Conclude that $\deg(\mathrm{GCD}(A, B)) = k$ iff $\mathbf{psc}_k(A, B) \neq 0$ and for all $i = 0, \ldots, k-1$, $\mathbf{psc}_i(A, B) = 0$.

5 Subresultant chains with indeterminate coefficients

We next follow Loos [BCL83] in developing the basic results of subresultants in terms of polynomials with indeterminate coefficients. In other words, for this section we assume that the input polynomials

$$A = \sum_{i=0}^{m} a_i x^i , \qquad B = \sum_{i=0}^{n} b_i x^i \tag{9}$$

come from the ring

$$\mathbf{Z}[x;\ a_m, \ldots, a_0, b_n, \ldots, b_0]$$

where a_i and b_i are indeterminates. After obtaining the properties of subresultants in this setting, we can interpret the indeterminate coefficients of A and B as values in I, thereby observing the behavior of the subresultant chain in $I[x]$. This approach has several advantages as noted by Loos [BCL83]. In particular it separates the sparsity effects (when consecutive chain members have degrees differing by more than one) from the similarity relations among various polynomials defined from the chain.

In this section, we assume that the polynomials A and B in (9) satisfies $m = n + 1$. We will see that this is without loss of generality in our applications.

Lemma 4 *Assume that the polynomials A and B in* (9) *satisfies $m = n + 1$. Then*

(i) $\mathbf{sres}_{n-1}(A, B) = \mathbf{prem}(A, B)$.

(ii) *For* $k = 0, \ldots, n-2$, $b_n^{2(n-k-1)} \mathbf{sres}_k(A, B) = \mathbf{sres}_k(B, \mathbf{prem}(A, B))$.

Proof. Part (i) is just a restatement of the previous lemma since

$$\begin{aligned} \mathbf{sres}_{n-1}(A, B) &= \mathbf{detpol}(A, xB, B) \\ &= \mathbf{detpol}(xB, B, A) = \mathbf{prem}(A, B). \end{aligned}$$

To see part (ii),

$$\begin{aligned}
\mathbf{sres}_k(A,B) &= \mathbf{detpol}(x^{n-1-k}A,\ldots,A,x^{n-k}B,\ldots,B)\\
&= \mathbf{detpol}(x^{n-1-k}b_n^2A,\ldots,b_n^2A,x^{n-k}B,\ldots,B)\cdot b_n^{-2(n-k)}\\
&= \mathbf{detpol}(x^{n-1-k}\mathbf{prem}(A,B),\ldots,\mathbf{prem}(A,B),x^{n-k}B,\ldots,B)\cdot b_n^{-2(n-k)}\\
&\qquad \text{(by subtracting appropriate multiples of B-rows from A-rows.)}\\
&= \mathbf{detpol}(x^{n-k}B,\ldots,B,x^{n-1-k}\mathbf{prem}(A,B),\ldots,\mathbf{prem}(A,B))\cdot b_n^{-2(n-k)}\\
&\qquad \text{(by transposing rows)}\\
&= \mathbf{detpol}(x^{n-k-2}B,\ldots,B,x^{n-1-k}\mathbf{prem}(A,B),\ldots,\mathbf{prem}(A,B))\cdot b_n^{-2(n-k-1)}\\
&\qquad \text{(using the fact that $k\le n-2$ and $\deg(\mathbf{prem}(A,B))=n-1$)}\\
&= \mathbf{sres}_k(B,\mathbf{prem}(A,B))\cdot b_n^{-2(n-k-1)}.
\end{aligned}$$

Q.E.D.

We note that Lemma 4 (i) justifies viewing subresultants as a generalization of pseudo-remainder.

Theorem 5 (*Habicht's theorem*) *Let $S_{n+1}, S_n, \ldots, S_0$ be the subresultant chain of A and B. Let R_j $(j = 0,\ldots,n+1)$ be the jth principal subresultant coefficient of A, B. Then for all $j = 1,\ldots,n$,*

$$R_{j+1}^{2(j-k)}S_k = \mathbf{sres}_k(S_{j+1},S_j)\,,\quad k=0,\ldots,j-1 \tag{10}$$

$$R_{j+1}^{2}S_{j-1} = \mathbf{prem}(S_{j+1},S_j). \tag{11}$$

Proof. We first observe that (11) follows from (10) since substituting $j-1$ for k in (10) gives $R_{j+1}^2S_{j-1} = \mathbf{sres}_{j-1}(S_{j+1},S_j) = \mathbf{prem}(S_{j+1},S_j)$ by the last lemma. It remains to show equation (10). Basis: If $j=n$ then equation (10) follows from the definitions. Induction: let $j<n$, assuming the lemma is true for $j+1$. Let $k\le j-1$. By the induction hypothesis

$$\begin{aligned}
R_{j+2}^{2(j-k+1)}S_k &= \mathbf{sres}_k(S_{j+2},S_{j+1})\\
&= R_{j+1}^{-2(j-k)}\mathbf{sres}_k(S_{j+1},\mathbf{prem}(S_{j+2},S_{j+1})) \quad \text{(by previous lemma (ii))}\\
&= R_{j+1}^{-2(j-k)}\mathbf{sres}_k(S_{j+1},R_{j+2}^2S_j) \quad \text{(by(11))}\\
&= R_{j+1}^{-2(j-k)}\mathbf{detpol}(x^{j-k-1}S_{j+1},\ldots,S_{j+1},R_{j+2}^2x^{j-k}S_j,\ldots,R_{j+2}^2S_j)\\
&= R_{j+1}^{-2(j-k)}R_{j+2}^{2(j-k+1)}\mathbf{sres}_k(S_{j+1},S_j).
\end{aligned}$$

This completes our proof of (10). **Q.E.D.**

6 Specialization of subresultant chains

We now want to substitute the indeterminate coefficients a_i, b_j in $\mathbf{Z}[x;\ a_{n+1}, \ldots, a_0, b_n, \ldots, b_0]$ with specific values in the domain I. This is called *specializing* the indeterminate coefficients and it induces a natural homomorphism from $\mathbf{Z}[x;\ a_{n+1}, \ldots, a_0, b_n, \ldots, b_0]$ to $I[x]$. Let σ denote any such homomorphism. Let S_k (resp. R_k) denote the kth subresultant (resp. principal subresultant coefficient) of A and B regarded as elements of $\mathbf{Z}[x;\ a_{n+1}, \ldots, a_0, b_n, \ldots, b_0]$. Let $\bar{A}, \bar{B}, \bar{S}_k$ and $\bar{R}_k$ denote the image of A, B, S_k and R_k under σ. In general, we write $\bar{a}$ over an expression to denote the image of that expression under σ, e.g. $\overline{\mathrm{sres}_i(S_{j+1}, S_j)}$. It is important to realize that this is not the same as $\mathrm{sres}_i(\bar{S}_{j+1}, \bar{S}_j)$. This fact seems to have been overlooked by Loos [BCL83] therefore rendering his proofs invalid. Our goal is to investigate the specialized chain

$$(\bar{S}_{n+1}, \ldots, \bar{S}_0).$$

Let d_k and k denote the actual and nominal degree of $\bar{S}_k$ respectively, with the usual convention that $\deg(0) = -\infty$. A subresultant $\bar{S}_k$ is said to be *defective of degree* r if $d_k = r < k$; otherwise $\bar{S}_k$ is said to be *regular*. A chain is said to be *defective* if any of its members is defective; otherwise it is *regular*. Thus for a defective chain, the sequence $(d_{n+1}, d_n, \ldots, d_0)$ contains gaps in the sense that some integer in the set $\{0, \ldots, n\}$ is missing. The growth of the content of polynomials in any PRS is closely linked to the presence of gaps. Note that $\mathrm{lead}(\bar{S}_k) \neq \bar{R}_k$ if $\bar{S}_k$ is defective or $k = n + 1$.

Definition. Let $A, B \in I[x]$ with $\deg(A) = m > \deg(B) = n > 0$. For any i, $0 \leq i < m - 1$, define the ith *pseudo subresultant* of A and B to be

$$\mathrm{psres}_i(A, B) = \mathrm{detpol}(x^{m-i-2}A, x^{m-i-3}A, \ldots, A, x^{m-i-1}B, x^{m-i-2}B, \ldots, B).$$

Note that if $\bar{S}_{j+1}$ is regular, then for $0 \leq k < j$

$$\mathrm{psres}_k(\bar{S}_{j+1}, \bar{S}_j) = \mathrm{detpol}(x^{j-1-k}\bar{S}_{j+1}, \ldots, \bar{S}_{j+1}, x^{j-k}\bar{S}_j, \ldots, \bar{S}_j) = \overline{\mathrm{sres}_k(S_{j+1}, S_j)}.$$

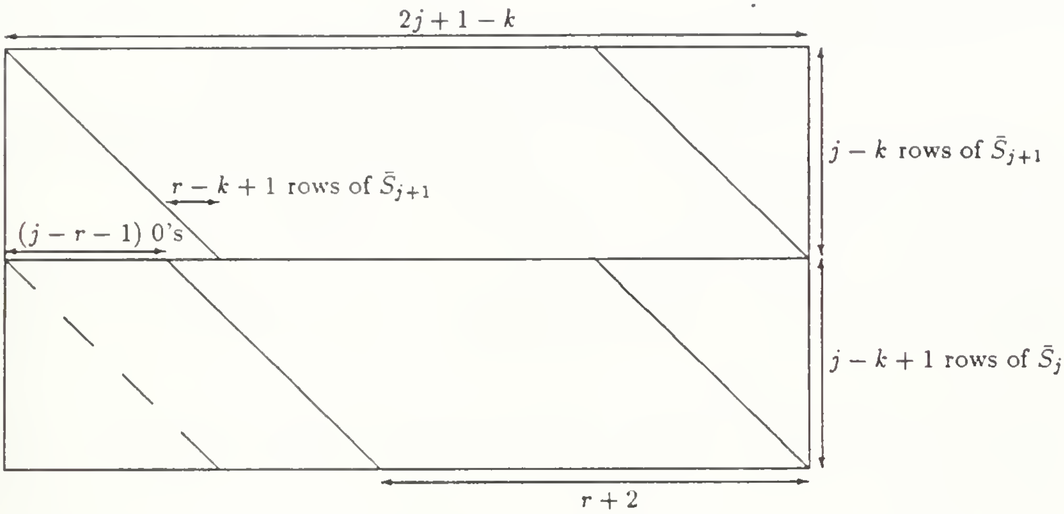

$2j+1-k$
$j-k$ rows of $\bar{S}_{j+1}$
$r-k+1$ rows of $\bar{S}_{j+1}$
$(j-r-1)$ 0's
$j-k+1$ rows of $\bar{S}_j$
$r+2$

Figure 1: The matrix associated to $\mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j)$

If both $\bar{S}_{n+1}$ and $\bar{S}_n$ are defective, then it is easily seen that the rest of the chain is 0. Henceforth, we assume that at least one of $\bar{S}_{n+1}$ or $\bar{S}_n$ will be regular. Without lost of generality, we may assume that $\bar{S}_{n+1}$ is regular. The following lemma is the specialization of Lemma 4:

Lemma 6 *Let $\bar{S}_{n+1}, \bar{S}_n, \ldots, \bar{S}_0$ be the pseudo subresultant chain of $\bar{A}$ and $\bar{B}$. Let ξ_j be the actual leading coefficient of $\bar{S}_j$. For $j = 1, \ldots, n$, if $\bar{S}_{j+1}$ is regular and $\bar{S}_j$ has degree $r+1$ $(-\infty \leq r+1 \leq j)$ then*

(i) $\mathbf{psres}_r(\bar{S}_{j+1}, \bar{S}_j) = \xi_{j+1}^{j-r-1}(-1)^{j-r+1}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)$.

(ii) $\mathbf{psres}_{r+1}(\bar{S}_{j+1}, \bar{S}_j) = \xi_{j+1}^{j-r-1}\xi_j^{(j-r-1)}\bar{S}_j$.

(iii) *For $k = 0, \ldots, r-1$,*

$$\begin{aligned} & \mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j) \\ = \; & \mathbf{psres}_k(\bar{S}_j, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \cdot \xi_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \end{aligned}$$

Proof. (i) By Figure 1, we have

$$\begin{aligned}\mathbf{psres}_r(\bar{S}_{j+1}, \bar{S}_j) &= \xi_{j+1}^{j-r-1}\mathbf{detpol}(\bar{S}_{j+1}, x^{j-r}\bar{S}_j, \ldots, \bar{S}_j) \\ &= \xi_{j+1}^{j-r-1}(-1)^{j-r+1}\mathbf{detpol}(x^{j-r}\bar{S}_j, \ldots, \bar{S}_j, \bar{S}_{j+1}) \\ &= \xi_{j+1}^{j-r-1}(-1)^{j-r+1}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j) \quad \text{(by Lemma 3)}\end{aligned}$$

Also from Figure 1, we can easily get (ii). To see part (iii),

$$\begin{aligned}&\mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j) \\ = \;& \mathbf{detpol}(x^{j-1-k}\bar{S}_{j+1}, \ldots, \bar{S}_{j+1}, x^{j-k}\bar{S}_j, \ldots, \bar{S}_j) \\ = \;& \mathbf{detpol}(x^{j-1-k}\bar{S}_{j+1}, \ldots, x^{r-k+1}\bar{S}_{j+1}, x^{r-k}\xi_j^{j-r+1}\bar{S}_{j+1}, \ldots, \xi_j^{j-r+1}\bar{S}_{j+1}, x^{j-k}\bar{S}_j, \ldots, \bar{S}_j) \\ & \cdot\xi_j^{-(j-r+1)(r+1-k)} \\ = \;& \mathbf{detpol}(x^{j-1-k}\bar{S}_{j+1}, \ldots, x^{r-k+1}\bar{S}_{j+1}, x^{r-k}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j), \ldots, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j), x^{j-k}\bar{S}_j, \ldots, \bar{S}_j) \\ & \cdot\xi_j^{-(j-r+1)(r+1-k)} \\ = \;& \mathbf{detpol}(x^{j-1-k}\bar{S}_{j+1}, \ldots, x^{r-k+1}\bar{S}_{j+1}, x^{j-k}\bar{S}_j, \ldots, x^{r-k}\bar{S}_j, x^{r-k-1}\bar{S}_j, \ldots, \bar{S}_j, x^{r-k}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j), \ldots, \\ & \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r+1-k)} \\ = \;& \mathbf{detpol}(x^{r-k-1}\bar{S}_j, \ldots, \bar{S}_j, x^{r-k}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j), \ldots, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \\ & \cdot\xi_{j+1}^{j-r-1} \cdot \xi_j^{j-r+1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r+1-k)} \\ = \;& \mathbf{psres}_k(\bar{S}_j, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \cdot \xi_{j+1}^{j-r-1} \cdot \xi_j^{j-r+1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r+1-k)}\end{aligned}$$

Q.E.D.

Theorem 7 (*Specialized Habicht's theorem*) *Let* $\bar{S}_{n+1}, \bar{S}_n, \ldots, \bar{S}_0$ *be the pseudo subresultant chain of* $\bar{A}$ *and* $\bar{B}$. *Let* $\bar{R}_j$ $(j = 0, \ldots, n-1)$ *be the nominal leading coefficient of* $\mathbf{psres}_j(\bar{A}, \bar{B})$. *Let* $\bar{R}_n = \mathbf{lead}(\bar{B})$ *and* $\bar{R}_{n+1} = 1$. *For* $j = 1, \ldots, n$, *if* $\bar{S}_{j+1}$ *is regular and* $\deg(\bar{S}_j) \leq j$ *then*

$$\bar{R}_{j+1}^{2(j-k)}\bar{S}_k = \mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j)\,, \quad k = 0, \ldots, j-1 \tag{12}$$

Proof. We prove the theorem by induction. Basis: If $j = n$ then equation (12) follows from the definitions. Induction: assuming the lemma is true for $j+1$ and $\bar{S}_j$ has degree $r+1$ $(-\infty \leq r+1 \leq j)$. We want to show that $\bar{R}_{r+1}^{2(r-k)}\bar{S}_k = \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r)$ for $k \leq r-1$. By the induction hypothesis $\bar{R}_{j+1}^{2(j-k)}\bar{S}_k = \mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j)$. We have two cases to consider:

Case (i) : $j = n$, then $\bar{R}_{j+1} = 1$. Thus, by the previous lemma (ii), we have

$$\bar{S}_{r+1} = \mathbf{psres}_{r+1}(\bar{S}_{j+1}, \bar{S}_j) = \xi_{j+1}^{(j-r-1)}\xi_j^{(j-r-1)}\bar{S}_j.$$

Hence,

$$\bar{S}_j = (\xi_{j+1}\xi_j)^{-(j-r-1)}\bar{S}_{r+1} \quad \text{and} \quad \bar{R}_{r+1} = \xi_{j+1}^{j-r-1}\xi_j^{j-r} \tag{13}$$

Also by the previous lemma (i), we have

$$\bar{S}_r = \mathbf{psres}_r(\bar{S}_{j+1}, \bar{S}_j) = \xi_{j+1}^{j-r-1}(-1)^{j-r+1}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j).$$

Thus,

$$(-\xi_{j+1})^{-(j-r-1)}\bar{S}_r = \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j) \tag{14}$$

Hence,

$$\begin{aligned}
&\bar{S}_k \\
= \ & \mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j) \quad \text{(by inductive hypothesis)} \\
= \ & \mathbf{psres}_k(\bar{S}_j, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \cdot \xi_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \\
= \ & \mathbf{detpol}(x^{r-k-1}\bar{S}_j, \ldots, \bar{S}_j, x^{r-k}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j), \ldots, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \\
& \cdot \xi_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \quad \text{(by previous lemma (iii))} \\
= \ & \mathbf{detpol}(x^{r-k-1}(\xi_{j+1}\xi_j)^{-(j-r-1)}\bar{S}_{r+1}, \ldots, (\xi_{j+1}\xi_j)^{-(j-r-1)}\bar{S}_{r+1}, x^{r-k}(-\xi_{j+1})^{-(j-r-1)}\bar{S}_r, \ldots, \\
& (-\xi_{j+1})^{-(j-r-1)}\bar{S}_r) \cdot \xi_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \quad \text{(by (13) and (14))} \\
= \ & \mathbf{detpol}(x^{r-k-1}\bar{S}_{r+1}, \ldots, \bar{S}_{r+1}, x^{r-k}\bar{S}_r, \ldots, \bar{S}_r) \cdot (\xi_{j+1} \cdot \xi_j)^{-(j-r-1)(r-k)} \cdot (-\xi_{j+1})^{-(j-r-1)(r-k+1)} \\
& \cdot \xi_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \xi_{j+1}^{-2(j-r-1)(r-k)} \cdot \xi_j^{-(j-r-1)(r-k)} \\
& \cdot (-1)^{(j-r+1)(r-k+1)} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \xi_{j+1}^{-2(j-r-1)(r-k)} \cdot (-1)^{(r+1-k)(2j-2r+2+r-k)} \cdot \xi_j^{-2(j-r)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \xi_{j+1}^{-2(j-r-1)(r-k)} \cdot \xi_j^{-2(j-r)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \bar{R}_{r+1}^{-2(r-k)} \quad \text{(by (13))}
\end{aligned}$$

Case (ii) : $j \neq n$, then $\xi_{j+1} = \bar{R}_{j+1}$. Thus, by the previous lemma (ii), we have

$$\bar{R}_{j+1}^{2(j-r-1)}\bar{S}_{r+1} = \mathbf{psres}_{r+1}(\bar{S}_{j+1}, \bar{S}_j) = \bar{R}_{j+1}^{j-r-1}\xi_j^{(j-r-1)}\bar{S}_j.$$

Hence,

$$\bar{S}_j = \bar{R}_{j+1}^{j-r-1}\xi_j^{-(j-r-1)}\bar{S}_{r+1} \quad \text{and} \quad \bar{R}_{r+1} = \bar{R}_{j+1}^{-(j-r-1)}\xi_j^{j-r} \tag{15}$$

Also by the previous lemma (i), we have

$$\bar{R}_{j+1}^{2(j-r)}\bar{S}_r = \mathbf{psres}_r(\bar{S}_{j+1}, \bar{S}_j) = \bar{R}_{j+1}^{j-r-1}(-1)^{j-r+1}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j).$$

Thus,

$$(-\bar{R}_{j+1})^{j-r+1}\bar{S}_r = \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j) \tag{16}$$

Hence,

$$\begin{aligned}
& \bar{R}_{j+1}^{2(j-k)}\bar{S}_k \\
= \ & \mathbf{psres}_k(\bar{S}_{j+1}, \bar{S}_j) \quad \text{(by inductive hypothesis)} \\
= \ & \mathbf{psres}_k(\bar{S}_j, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \cdot \xi_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \\
= \ & \mathbf{detpol}(x^{r-k-1}\bar{S}_j, \ldots, \bar{S}_j, x^{r-k}\mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j), \ldots, \mathbf{prem}(\bar{S}_{j+1}, \bar{S}_j)) \\
& \cdot \bar{R}_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \quad \text{(by previous lemma (iii))} \\
= \ & \mathbf{detpol}(x^{r-k-1}\bar{R}_{j+1}^{j-r-1}\xi_j^{-(j-r-1)}\bar{S}_{r+1}, \ldots, \bar{R}_{j+1}^{j-r-1}\xi_j^{-(j-r-1)}\bar{S}_{r+1}, x^{r-k}(-\bar{R}_{j+1})^{j-r+1}\bar{S}_r, \ldots, \\
& (-\bar{R}_{j+1})^{j-r+1}\bar{S}_r) \cdot \bar{R}_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \quad \text{(by (15) and (16))} \\
= \ & \mathbf{detpol}(x^{r-k-1}\bar{S}_{r+1}, \ldots, \bar{S}_{r+1}, x^{r-k}\bar{S}_r, \ldots, \bar{S}_r) \cdot \bar{R}_{j+1}^{(j-r-1)(r-k)} \cdot \xi_j^{-(j-r-1)(r-k)} \cdot (-\bar{R}_{j+1})^{(j-r+1)(r-k+1)} \\
& \cdot \bar{R}_{j+1}^{j-r-1} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \bar{R}_{j+1}^{(j-r-1)(r-k+1)} \cdot \xi_j^{-(j-r-1)(r-k)} \cdot \bar{R}_{j+1}^{(j-r+1)(r-k+1)} \\
& \cdot (-1)^{(j-r+1)(r-k+1)} \cdot (-1)^{(r+1-k)(j-k+1)} \cdot \xi_j^{-(j-r+1)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \bar{R}_{j+1}^{2(j-r)(r-k+1)} \cdot (-1)^{(r+1-k)(2j-2r+2+r-k)} \cdot \xi_j^{-2(j-r)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \bar{R}_{j+1}^{2(j-r-1)(r-k)} \cdot \bar{R}_{j+1}^{2(j-k)} \cdot \xi_j^{-2(j-r)(r-k)} \\
= \ & \mathbf{psres}_k(\bar{S}_{r+1}, \bar{S}_r) \cdot \bar{R}_{r+1}^{-2(r-k)} \cdot \bar{R}_{j+1}^{2(j-k)} \quad \text{(by (15))}
\end{aligned}$$

This completes our proof of (12). **Q.E.D.**

Theorem 8 (*Pseudo Subresultant Theorem*) *Let* $(\bar{S}_{n+1}, \ldots, \bar{S}_0)$ *be a specialized chain in* $I[x]$. *For* $j = 1, \ldots, n$, *if* $\bar{S}_{j+1}$ *is regular and* $\bar{S}_j$ *has degree* r $(-\infty \leq r \leq j)$ *then*

1. $j = n$,

$$\bar{S}_{j-1} = \bar{S}_{j-2} = \cdots = \bar{S}_{r+1} = 0 \tag{17}$$

$$\bar{S}_r = \mathbf{lead}(\bar{S}_j)^{j-r}\mathbf{lead}(\bar{S}_{j+1})^{j-r}\bar{S}_j, \quad 0 \leq r \leq j \tag{18}$$

$$\bar{S}_{r-1} = (-\mathbf{lead}(\bar{S}_{j+1}))^{j-r}\mathbf{prem}(\bar{S}_{j+1},\bar{S}_j), \quad 1 \le r \le j \tag{19}$$

2. $j < n$,

$$\bar{S}_{j-1} = \bar{S}_{j-2} = \cdots = \bar{S}_{r+1} = 0 \tag{20}$$

$$(\bar{R}_{j+1})^{j-r}\bar{S}_r = \mathbf{lead}(\bar{S}_j)^{j-r}\bar{S}_j, \quad 0 \le r \le j \tag{21}$$

$$(-\bar{R}_{j+1})^{j-r+2}\bar{S}_{r-1} = \mathbf{prem}(\bar{S}_{j+1},\bar{S}_j), \quad 1 \le r \le j \tag{22}$$

Proof. From (12) we see that for $k = r-1, r, \ldots, j-1$,

$$\bar{R}_{j+1}^{2(j-k)}\bar{S}_k = \mathbf{psres}_k(\bar{S}_{j+1},\bar{S}_j). \tag{23}$$

The right-hand side of (23) is equal to $\mathbf{detpol}(M)$ where M has the form as shown in Figure 1. (Note: In Figure 1, the degree of $\bar{S}_j$ is $r+1$.) Observe that $\bar{S}_j$ has $j-r$ leading terms vanishing, and there are $j-k$ rows of $\bar{S}_{j+1}$ in the matrix M. If $j-r \ge j-k$, i.e. $k \ge r$, then it is clear from the picture of M that M will be upper diagonal (i.e. has only zeroes below its main diagonal). Observe that the coefficients in $\mathbf{detpol}(M)$ are all determinants of upper triangular square matrices. We consider three cases.

Case $k = r$ and $j = n$: Then $\bar{R}_{j+1} = 1$ and $\mathbf{detpol}(M)$ evaluates to $\mathbf{lead}(\bar{S}_{j+1})^{j-r}\mathbf{lead}(\bar{S}_j)^{j-r}S_j$. This shows (18).

Case $k = r$ and $j < n$: Then $\mathbf{detpol}(M)$ evaluates to $\bar{R}_{j+1}^{j-r}\mathbf{lead}(\bar{S}_j)^{j-r}S_j$. This shows (21).

Case $k > r$: Each diagonal of the square matrices determining coefficients of $\mathbf{detpol}(M)$ has at least one zero element, hence each coefficient is zero. This shows (17) and (20).

Now suppose $j-r = j-k-1$, i.e. $k = r-1$. Then, on expanding the first $j-r$ rows of $\mathbf{detpol}(M)$, we obtain

$$\begin{aligned}
\mathbf{detpol}(M) &= \mathbf{lead}(\bar{S}_{j+1})^{j-r}\mathbf{detpol}(\bar{S}_{j+1}, x^{j-r+1}\bar{S}_j, x^{j-r}\bar{S}_j, \ldots, \bar{S}_j) \\
&= (-\mathbf{lead}(\bar{S}_{j+1}))^{j-r}\mathbf{detpol}(x^{j-r+1}\bar{S}_j, \ldots, \bar{S}_j, \bar{S}_{j+1}) \\
&= (-\mathbf{lead}(\bar{S}_{j+1}))^{j-r}\mathbf{prem}(\bar{S}_{j+1},\bar{S}_j).
\end{aligned} \tag{24}$$

If $j = n$, then $\bar{R}_{j+1} = 1$. This, combined with (23) and (24), implies (19). If $j < n$, then $\bar{R}_{j+1} = \mathbf{lead}(\bar{S}_{j+1})$. This, combined with (23) and (24), implies (22). **Q.E.D.**

Let $(\bar{S}_{n+1}, \bar{S}_n, \ldots, \bar{S}_0)$ be the pseudo subresultant chain of $\bar{A}$ and $\bar{B}$. Let $(S_{n+1}, \ldots, S_0)$ be the subresultant chain of $\bar{A}$ and $\bar{B}$. Let R_j $(j = 0, \ldots, n+1)$ be the jth principal subresultant coefficient of $\bar{A}, \bar{B}$. Suppose the degree of $\bar{B}$ is $l < n + 1$. Then clearly

$$\bar{S}_j = \mathbf{lead}(S_{n+1})^{n-l} S_j \quad \text{and} \quad \bar{R}_j = \mathbf{lead}(S_{n+1})^{n-l} R_j \quad \text{for} \quad 0 \leq j < n. \tag{25}$$

Before showing the subresultant structure theorem, we note the following useful consequence of lemma 3:

Corollary 9 *Let $A, B \in I[x]$ and $\alpha, \beta \in I$. Then*

$$\mathbf{prem}(\alpha A, \beta B) = \alpha \beta^{j-r+2} \mathbf{prem}(A, B)$$

where $j + 1 = \deg(A) \geq r = \deg(B)$.

Theorem 10 (*Subresultant Structure Theorem*) *Let $(S_{n+1}, \ldots, S_0)$ be the subresultant chain of $\bar{A}$ and $\bar{B}$. For $j = 1, \ldots, n$, if S_{j+1} is regular and S_j has degree r $(-\infty \leq r \leq j)$ then*

$$S_{j-1} = S_{j-2} = \cdots = S_{r+1} = 0 \tag{26}$$

$$(R_{j+1})^{j-r} S_r = \mathbf{lead}(S_j)^{j-r} S_j, \quad r \geq 0 \tag{27}$$

$$(-R_{j+1})^{j-r+2} S_{r-1} = \mathbf{prem}(S_{j+1}, S_j), \quad r \geq 1 \tag{28}$$

Proof. From (20) and (25), we immediately get (26). To see (27), we consider the following two cases. Case one: if $j = n$, by definition $R_{j+1} = 1$ and $\bar{S}_{j+1} = S_{j+1}$ and $\bar{S}_j = S_j$. Thus, the right hand side of (18) is equal to $\mathbf{lead}(S_{j+1})^{j-r} \mathbf{lead}(S_j)^{j-r} S_j$ and by (25) the left hand side of (18) is equal to $\textbf{lead}(S_{j+1})^{j-r}S_r = \textbf{lead}(S_{j+1})^{j-r}R_{j+1}^{j-r}S_r$. Hence,

$$R_{j+1}^{j-r}S_r = \textbf{lead}(S_j)^{j-r}S_j \quad \text{for } j = n, r \geq 0 \tag{29}$$

Case two: if $j < n$, then by (25), (21) can be evaluated to

$$(\textbf{lead}(S_{n+1})^{n-l}R_{j+1})^{j-r}\textbf{lead}(S_{n+1})^{n-l}S_r = \textbf{lead}(\textbf{lead}(S_{n+1})^{n-l}S_j)^{j-r}\textbf{lead}(S_{n+1})^{n-l}S_j$$

(Note that this equation is true even when $j+1 = n$ because in this case $\bar{S}_{j+1}$ is regular and then $n = l$ and thus $\textbf{lead}(S_{n+1})^{n-l} = 1$ so that we can add it anywhere in this equation.) Hence, we have when $j < n$

$$(R_{j+1})^{j-r}S_r = \textbf{lead}(S_j)^{j-r}S_j, \quad \text{for } j < n, r \geq 0$$

This with (29) shows (27).

To see (28), we also consider the following two cases.

Case one: if $j = n$, by definition $R_{j+1} = 1$ and $\bar{S}_{j+1} = S_{j+1}$ and $\bar{S}_j = S_j$. Thus, the right hand side of (19) is equal to $(-\textbf{lead}(S_{j+1}))^{j-r}\textbf{prem}(S_{j+1}, S_j)$ and by (25) the left hand side of (19) is equal to $\textbf{lead}(S_{j+1})^{j-r}S_{r-1} = \textbf{lead}(S_{j+1})^{j-r}R_{j+1}^{j-r+2}S_{r-1}$. Hence,

$$(-R_{j+1})^{j-r+2}S_{r-1} = \textbf{prem}(S_{j+1}, S_j), \quad \text{for } j = n, r \geq 1 \tag{30}$$

Case two: if $j < n$, then by (25) and Corollary 9, (22) can be evaluated to

$$\begin{aligned} &(-\textbf{lead}(S_{n+1})^{n-l}R_{j+1})^{j-r+2}\textbf{lead}(S_{n+1})^{n-l}S_{r-1} \\ &= \textbf{prem}(\textbf{lead}(S_{n+1})^{n-l}S_{j+1}, \textbf{lead}(S_{n+1})^{n-l}S_j) \\ &= \textbf{lead}(S_{n+1})^{n-l}(\textbf{lead}(S_{n+1})^{n-l})^{j-r+2}\textbf{prem}(S_{j+1}, S_j) \end{aligned}$$

(Note that the above equation is true even when $j+1 = n$.) Hence, we have when $j < n$

$$(-R_{j+1})^{j-r+2}S_{r-1} = \textbf{prem}(S_{j+1}, S_j), \quad \text{for } j < n, r \geq 1$$

This with (30) shows (28). **Q.E.D.**

Note: This theorem is essentially the "Subresultant Theorem" in Loos [BCL83] (page 122). The subresultant structure theorem describes the gap structure in subresultant chains. In particular, it says we can uniquely divide the chain into *blocks* of (consecutive) subresultants such that if

$$(S_i, S_{i-1}, \ldots, S_{j+1}, S_j)\,, \quad n+1 \geq i \geq j \geq 0,$$

is a block then

1. Either $j = 0$ and all the subresultants in the block are zero, or else both S_i and S_j are non-zero, $S_i \sim S_j$ and $S_{i-1} = S_{i-2} = \cdots = S_{j+1} = 0$. Note that there is at most one block where all the subresultants are zero; call this the *zero block*.

2. Assume that the block under consideration is a non-zero block. Then S_j is regular. Hence if $i > j$ then S_i is defective. Thus every defective subresultant S_i of degree r corresponds to a unique regular subresultant S_j, $r = j$, that is similar to it. This also implies that there cannot be two consecutive non-zero subresultants that are both defective.

7 Relation between Subresultant Chain and Subresultant PRS

Now we relate the subresultant PRS $\mathbf{p}$ which is described in Section 3 (equations (5) and (6)) to the subresultant chain

$$\mathbf{S} = (S_{n+1}, S_n, \ldots, S_0)$$

where $S_{n+1} = p_0$ and $S_n = p_1$. (Note that the convention for subscripting for PRS's in increasing order is opposite to that for subresultant chains.) By the definition of PRS, $\deg(p_0) \geq \deg(p_1)$. In Section 6, we only discuss the case $\deg(p_0) > \deg(p_1)$. Now, we need to consider the case $\deg(p_0) = \deg(p_1)$.

Suppose $A, B \in I[x]$ and $\deg(A) = \deg(B) = n$. Let $(S_{n+1}, S_n, \ldots, S_0)$ be the subresultant chain of A and B. Let $(R_{n+1}, R_n, \ldots, R_0)$ be the principal subresultant coefficients of A and B. Then

Lemma 11 $S_{n-1} = -\mathbf{prem}(A, B)$.

Proof.

$$\begin{aligned} S_{n-1} &= \mathbf{sres}_{n-1}(A, B) \\ &= \mathbf{detpol}(A, B) \quad = \quad -\mathbf{detpol}(B, A) \quad = \quad -\mathbf{prem}(A, B) \end{aligned}$$

Q.E.D.

Let $\mathbf{deg}(S_{n-1}) = l$ and $\mathbf{lead}(S_n) = \xi$.

Lemma 12 *For* $i = 0, \ldots, l$, $S_i = \xi^{i-l}\mathbf{sres}_i(S_n, S_{n-1})$.

Proof.

$$\begin{aligned} S_i &= \mathbf{sres}_i(S_{n+1}, S_n) \\ &= \mathbf{detpol}(x^{n-i-1}S_{n+1}, \ldots, S_{n+1}, x^{n-i-1}S_n, \ldots, S_n) \\ &= \mathbf{detpol}(x^{n-i-1}\xi S_{n+1}, \ldots, \xi S_{n+1}, x^{n-i-1}S_n, \ldots, S_n) \cdot \xi^{i-n} \\ &= \mathbf{detpol}(x^{n-i-1}\mathbf{prem}(S_{n+1}, S_n), \ldots, \mathbf{prem}(S_{n+1}, S_n), x^{n-i-1}S_n, \ldots, S_n) \cdot \xi^{i-n} \\ &= \mathbf{detpol}(x^{n-i-1}S_{n-1}, \ldots, S_{n-1}, x^{n-i-1}S_n, \ldots, S_n) \cdot (-\xi)^{i-n} \\ &= \mathbf{detpol}(x^{n-i-1}S_n, \ldots, S_n, x^{n-i-1}S_{n-1}, \ldots, S_{n-1}) \cdot (-\xi)^{i-n} \cdot (-1)^{(n-i)(n-i)} \\ &= \mathbf{detpol}(x^{n-i-1}S_n, \ldots, x^{l-i}S_n, x^{l-i-1}S_n, \ldots, S_n, x^{n-i-1}S_{n-1}, \ldots, S_{n-1}) \cdot \xi^{i-n} \\ &= \mathbf{sres}_i(S_n, S_{n-1}) \cdot \xi^{n-l} \cdot \xi^{i-n} \\ &= \xi^{i-l}\mathbf{sres}_i(S_n, S_{n-1}) \end{aligned}$$

Q.E.D.

Let $(P_n, P_{n-1}, \ldots, P_0)$ be the subresultant chain of S_n and S_{n-1}. Let $(C_n, C_{n-1}, \ldots, C_0)$ be the principal subresultant coefficients of S_n and S_{n-1}. Then, from Lemma 11, for $i = 0, \ldots, l$,

$$S_i = \xi^{i-l}P_i \text{ and } R_i = \xi^{i-l}C_i. \tag{31}$$

Theorem 13 *Suppose* $A, B \in I[x]$ *and* $\mathbf{deg}(A) = \mathbf{deg}(B) = n$. *Let* $(S_{n+1}, \ldots, S_0)$ *be the subresultant chain of* A *and* B. *Suppose* $\mathbf{deg}(S_{n-1}) = l$. *Then,*

$$S_l = \mathbf{lead}(S_{n-1})^{n-1-l}S_{n-1} \tag{32}$$

$$(-1)^{n-l+1}\xi S_{l-1} = \mathbf{prem}(S_n, S_{n-1}) \tag{33}$$

For $j = 1, \ldots, n-2$, if S_{j+1} is regular and S_j has degree r $(-\infty \le r \le j)$ then

$$S_{j-1} = S_{j-2} = \cdots = S_{r+1} = 0 \tag{34}$$

$$(R_{j+1})^{j-r} S_r = \mathbf{lead}(S_j)^{j-r} S_j, \quad r \ge 0 \tag{35}$$

$$(-R_{j+1})^{j-r+2} S_{r-1} = \mathbf{prem}(S_{j+1}, S_j), \quad r \ge 1 \tag{36}$$

Proof. Because $\mathbf{deg}(S_n) > \mathbf{deg}(S_{n-1})$, we can apply Theorem 10 so that for $j = 1, \ldots, n-1$, if P_{j+1} is regular and P_j has degree r $(-\infty \le r \le j)$ then

$$P_{j-1} = P_{j-2} = \cdots = P_{r+1} = 0$$

$$(C_{j+1})^{j-r} P_r = \mathbf{lead}(P_j)^{j-r} P_j, \quad r \ge 0$$

$$(-C_{j+1})^{j-r+2} P_{r-1} = \mathbf{prem}(P_{j+1}, P_j), \quad r \ge 1$$

Thus, $P_l = \mathbf{lead}(P_{n-1})^{n-1-l} P_{n-1}$. Hence, by (31), we have $S_l = \mathbf{lead}(S_{n-1})^{n-1-l} S_{n-1}$ which proves (32). Since $(-1)^{n-l+1} P_{l-1} = \mathbf{prem}(P_n, P_{n-1})$, we have, by (31), $(-1)^{n-l+1}\xi S_{l-1} = \mathbf{prem}(S_n, S_{n-1})$ which proves (33).

Also by (31), we have for $j = 1, \ldots, l-1$, if S_{j+1} is regular and S_j has degree r $(-\infty \le r \le j)$ then

$$\xi^{l-(j-1)} S_{j-1} = \xi^{l-(j-2)} S_{j-2} = \cdots = \xi^{l-(r+1)} S_{r+1} = 0$$

$$(\xi^{l-(j+1)} R_{j+1})^{j-r} \xi^{l-r} S_r = \mathbf{lead}(\xi^{l-j} S_j)^{j-r} \xi^{l-j} S_j, \quad r \ge 0$$

$$(-\xi^{l-(j+1)} R_{j+1})^{j-r+2} \xi^{l-(r-1)} S_{r-1} = \mathbf{prem}(\xi^{l-(j+1)} S_{j+1}, \xi^{l-j} S_j), \quad r \ge 1$$

But ξ in these equations can be canceled, thus we get (34) (35) and (36). **Q.E.D.**

Recall the block structure of subresultant chains. We make the following observations:

Lemma 14 *1. The number of non-zero blocks in the chain* **S** *is equal to the length $h+1$ of the subresultant PRS* p. *Hence we can write* **S**, *omitting its zero block if any, in the form*

$$\mathbf{S} = \mathbf{S}_0;\ \mathbf{S}_1;\ \ldots;\ \mathbf{S}_h$$

where $\mathbf{S}_k$ *is the kth block.*

2. For each k, let $S_{d(k)}$ denotes the regular subresultant, clearly $0 \le d(k) \le n+1$, in the kth block $\mathbf{S}_k$ *is similar to p_k in the subresultant PRS.*

Proof. We show by induction on k that the ith polynomial in p is similar to $S_{d(i)}$ for $i = 0, \ldots, h$. The result is true for $k = 1$ since $h \ge 2$ implies p_0 and p_1 are in separate blocks. If $k \ge 2$,

$$\begin{aligned} p_k &= \mathrm{prem}(p_{k-2}, p_{k-1}) \\ &\sim \mathrm{prem}(S_{d(k-2)}, S_{d(k-1)}) && \text{(by induction hypothesis)} \\ &\sim S_{d(k)} && \text{(from the subresultant structure theorem)} \end{aligned}$$

Q.E.D.

We now want to find the coefficients of similarity between $S_{d(k)}$ and p_k, $k = 0, \ldots, h$. By definition, the first subresultant in the block $\mathbf{S}_k$ is $S_{d(k-1)-1}$ (where we take $d(-1) = n+2$). It is convenient to let $e(k) = d(k-1) - 1$. Call $(S_{e(0)}, \ldots, S_{e(h)})$ the *subregular subresultants.* Note that all the nonzero defective subresultants are included among the subregular subresultants.

Theorem 15 (*Correctness of the subresultant PRS*) *The subregular subresultants in* **S** *is related to the subresultant PRS* p *by*

(i) $R_{d(i)} = \psi_i$, $i = 0, 1, \ldots, h$. *(except that $\psi_1 = 1$ when* $\mathbf{deg}(p_0) = \mathbf{deg}(p_1)$.)

(ii) $S_{e(i)} = p_i$, $i = 0, 1, \ldots, h$.

Proof. We proof both parts by induction on i.

Basis: If $i = 0$ then $R_{d(0)} = R_{n+1} = 1 = \psi_0$ by the definitions. Also, by definition, $S_{e(1)} = S_n = p_1$ and $S_{e(0)} = S_{d(0)} = S_{n+1} = p_0$. Thus

$$p_2 = \frac{\mathrm{prem}(p_0, p_1)}{\beta_1} = \frac{\mathrm{prem}(S_{d(0)}, S_{e(1)})}{\beta_1} = \frac{(-1)^{\delta_0+1} S_{e(2)}}{(-1)^{\delta_0+1}} = S_{e(2)}$$

If $\deg(p_0) = \deg(p_1)$, then $\delta_0 = 0$, so that $\psi_1 = 1$ and $\psi_2 = c_2^{\delta_1}$. From (32), we know that

$$R_{d(2)} = \mathbf{lead}(S_l) = c_2^{\delta_1} = \psi_2.$$

And from (33), we have

$$(-1)^{n-l+1}\xi S_{e(3)} = \mathbf{prem}(S_{e(1)}, S_{e(2)})$$

Hence,

$$S_{e(3)} = \frac{\mathbf{prem}(S_{e(1)}, S_{e(2)})}{(-1)^{n-l+1}c_1} = p_3$$

Inductive steps: for $i \geq 1$, assuming that part (i) is true for $i-1$ and part (ii) is true for i and $i+1$, we want to show that (i) hold for i and (ii) hold for $i+2$. We rewrite (27) by using the notation $d(i)$ and $e(i)$ as follows:

$$(R_{d(i-1)})^{\delta_{i-1}-1} S_{d(i)} = \mathbf{lead}(S_{e(i)})^{\delta_{i-1}-1} S_{e(i)} \tag{37}$$

Thus $(\psi_{i-1})^{\delta_{i-1}-1} S_{d(i)} = \mathbf{lead}(p_i)^{\delta_{i-1}-1} p_i$ and $(\psi_{i-1})^{\delta_{i-1}-1} R_{d(i)} = c_i^{\delta_{i-1}}$. Hence,

$$R_{d(i)} = \frac{c_i^{\delta_{i-1}}}{\psi_{i-1}^{\delta_{i-1}-1}} = \psi_i$$

which proves part (i). To see part (ii), we rewrite (28) as follows:

$$(-R_{d(i)})^{\delta_i+1} S_{e(i+2)} = \mathbf{prem}(S_{d(i)}, S_{e(i+1)}) \tag{38}$$

Then

$$\begin{aligned}
&\beta_{i+1} p_{i+2} = \mathbf{prem}(p_i, p_{i+1}) \\
&= \mathbf{prem}(S_{e(i)}, S_{e(i+1)}) \quad \text{(by inductive hypothesis.)} \\
&= \mathbf{prem}\Big(\frac{(R_{d(i-1)})^{\delta_{i-1}-1}}{\mathbf{lead}(S_{e(i)})^{\delta_{i-1}-1}} S_{d(i)}, S_{e(i+1)}\Big) \quad \text{(by (37))} \\
&= \frac{\psi_{i-1}^{\delta_{i-1}-1}}{c_i^{\delta_{i-1}-1}} \mathbf{prem}(S_{d(i)}, S_{e(i+1)}) \quad \text{(by inductive hypothesis and Corollary 9.)} \\
&= \frac{\psi_{i-1}^{\delta_{i-1}-1}}{c_i^{\delta_{i-1}-1}} (-\psi_i)^{\delta_i+1} S_{e(i+2)} \quad \text{(by (38) and part (i))} \\
&= \beta_{i+1} S_{e(i+2)}
\end{aligned}$$

Hence, $S_{e(i+2)} = p_{i+2}$, which proves part (ii). **Q.E.D**

Corollary 16 *The ψ_i's defined in (6) are elements of the integral domain I.*

References

[B71] Brown, W.S.: On Euclid's Algorithm and the Computation of Polynomial Greatest Common Division. *JACM 18, 476-504, 1971.*

[BT71] Brown, W.S., and Traub, J.F.: On Euclid's Algorithm and the Theory of Subresultants. *JACM 18, 505-514, 1971.*

[BCL83] Buchberger, B., Collins, G.E., and Loos, R.: Computer Algebra, 2nd ed., Spring-Verlag, 1983. *115-137.*

[C66] Collins, G.E.: Polynomial Remainder Sequences and Determinants. *Am. Math. Monthly 73, 708-712, 1966.*

[C67] Collins, G.E.: Subresultants and Reduced Polynomial Remainder Sequence. *JACM 14, 128-142, 1967.*

[SS83] Schwartz, J., and Scharir, M.: On the piano movers' problem: II. General techniques for computing topological properties of real algebraic manifolds. *Advances in Applied Math. 4, 298-351, 1983.*

[Y85] Yap, C.K.: Lecture Notes on Symbolic Computation, Class notes, Courant Institute, NYU, Fall 1985.

NYU COMPSCI TR-319 c.1
Ho, Chung-jen
Polynomial remainder
sequences and theory of
subresultants

NYU COMPSCI TR-319 c.1
Ho, Chung-jen
Polynomial remainder
sequences and theory of
subresultants

This book may be kept

FOURTEEN DAYS

A fine will be charged for each day the book is kept overtime.

GAYLORD 142

PRINTED IN U.S.A.